ESTIMATING MEAN LIFE FROM LIMITED TESTING

Donald W. Rankin
Army Materiel Test and Evaluation Directorate
US Army White Sands Missile Range
White Sands Missile Range, New Mexico 88002

ABSTRACT. Exact probability formulae are developed, with no restrictive assumptions, for use with tests which produce data of the constant failure rate type. Although universally valid, the formulae are particularly apropos when straitened test circumstances are dictated. Programming suggestions are included.

1. INTRODUCTION. This paper is a sequel to the one entitled *Estimating Reliability from Small Samples* and presented before the twenty-second conference on the Design of Experiments in October 1976 [4].

The Poisson distribution is treated in a manner parallel to that afforded the binomial distribution in the earlier paper.

2. DEFINITION OF *EVENT*. Probability statistics require the identification of a unit commonly called *event* or *trial*. Often this identification is self-evident. Suppose a test consists of drawing a sample of specified size (n, say) from a larger population of similar items, then determining the number of defective items (k) in the sample. It requires no stretch of the imagination to say that drawing that sample of size n constitutes an *event* or *trial* and that the failure ratio k/n is the *result* of that event. It is to be noted that the failure ratio is dimensionless; i.e., k and n are measured in the same units.

Identification is not always so clear-cut. For example, suppose an operator of heavy trucks notices that in the preceding six months, he has experienced 13 major mechanical breakdowns--one every two weeks, on the average. The definition of *failure* is obvious, but what is a *success*? To what do we add k to get n, the sample size? The mathematical answer is that $n \to \infty$. But this is also a useless answer; no realistic test design could require an infinite sample size.

To avoid facing this dilemma, let us arbitrarily define *event* in some convenient unit different from that in which k is expressed. As a consequence, we no longer have a failure ratio. In its place we substitute a failure rate--of k per event. Thus the failure rate depends upon an observed k, but upon a defined *event*.

To return to the truck operator, let us say that examination of the log books reveals a total operating mileage of 267150 for the period in question. This figure (267150 miles) is taken as the definition of *event*. The observed failure rate then becomes

$$\frac{13 \text{ failures}}{267150 \text{ miles}} = 0.0000486667 \text{ failures per mile.}$$

It is sometimes regarded as preferable to express the reciprocal of the failure rate, calling it *mean life*. Thus we would have

$$\frac{267150 \text{ miles}}{13 \text{ failures}} = 20550 \text{ mean miles between failures.}$$

The term *event* can be defined in any of a variety of units--area, volume, weight, time--almost anything that can be measured.

3. POISSON PROBABILITY. Consider the well-known series

$$e^{x} = \sum_{k=0}^{\infty} \frac{x^{k}}{\underline{|k}}, \ (k = 0, 1, 2, 3, \ldots) \tag{1}$$

This series converges for all finite values of x, provided only that x remains constant. Multiplying by e^{-x} produces

$$1 = \sum_{k=0}^{\infty} \frac{x^{k} e^{-x}}{\underline{|k}} \tag{2}$$

Poisson noted (1837) that if x is a constant failure rate and k is a non-negative integer, the probability of observing exactly k failures during an event is given by the appropriate term of the above expansion; i.e., by

$$p(k) = \frac{x^{k} e^{-x}}{\underline{|k}} \tag{3}$$

This last expression, then is a probability function in the discrete variable k. Unfortunately, however, it does not suffice. In most test designs, it will be possible to define *event* arbitrarily and to observe the value of k exactly, but nothing will be known about x. Usually, in fact, x will be the principal value sought. A probability function in x is required.

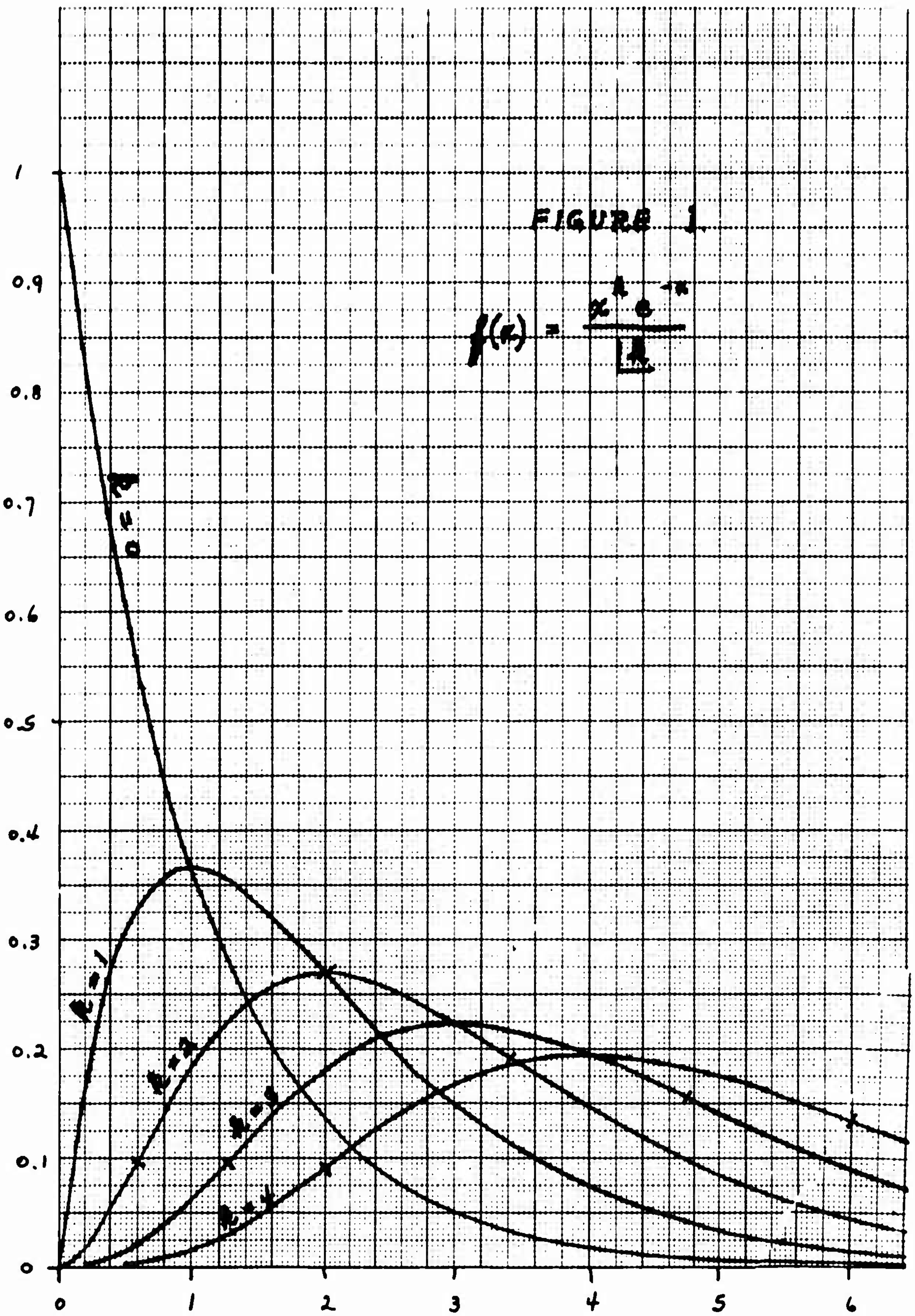


FIGURE 1
$f(x) = \frac{x^k e^{-x}}{k!}$
k = 0
k = 1
k = 2
k = 3
k = 4
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0
1
2
3
4
5
6

Now x can take on any non-negative value; i.e., it is a <u>continuous</u> variable within the limits $0 \leq x < \infty$. Necessarily

$$1 = \int_0^\infty f(x)\, dx$$

defines $f(x)$ as the required probability function in x, whatever form it may take. With k fixed, the expression

$$\frac{x^k e^{-x}}{\underline{|k}}$$

becomes a density function in x (though not necessarily a probability function). It is necessary to evaluate the definite integral

$$I_k = \int_0^\infty \frac{x^k e^{-x}}{\underline{|k}}\, dx$$

Since k is constant, $\underline{|k}$ can be taken outside the integral sign, leaving

$$\underline{|k}\, I_k = \int_0^\infty x^k e^{-x}\, dx = \Gamma\,(k + 1)$$

but also, k is an integer, hence $\underline{|k} = \Gamma\,(k + 1)$.

It is seen that $I_k = 1$, and therefore that

$$f(x) = \frac{x^k e^{-x}}{\underline{|k}}$$

is the required probability function in the continuous variable x.

It is helpful to inspect graphs of the probability function

$$f(x) = \frac{x^k e^{-x}}{\underline{|k}} \tag{4}$$

Several are depicted in Figure 1 for various integer values of k. Among features which should be noted are the following:

1) When $k = 0$, the function degenerates to

$$f(x) = e^{-x} \tag{5}$$

and is most easily treated as a separate case.

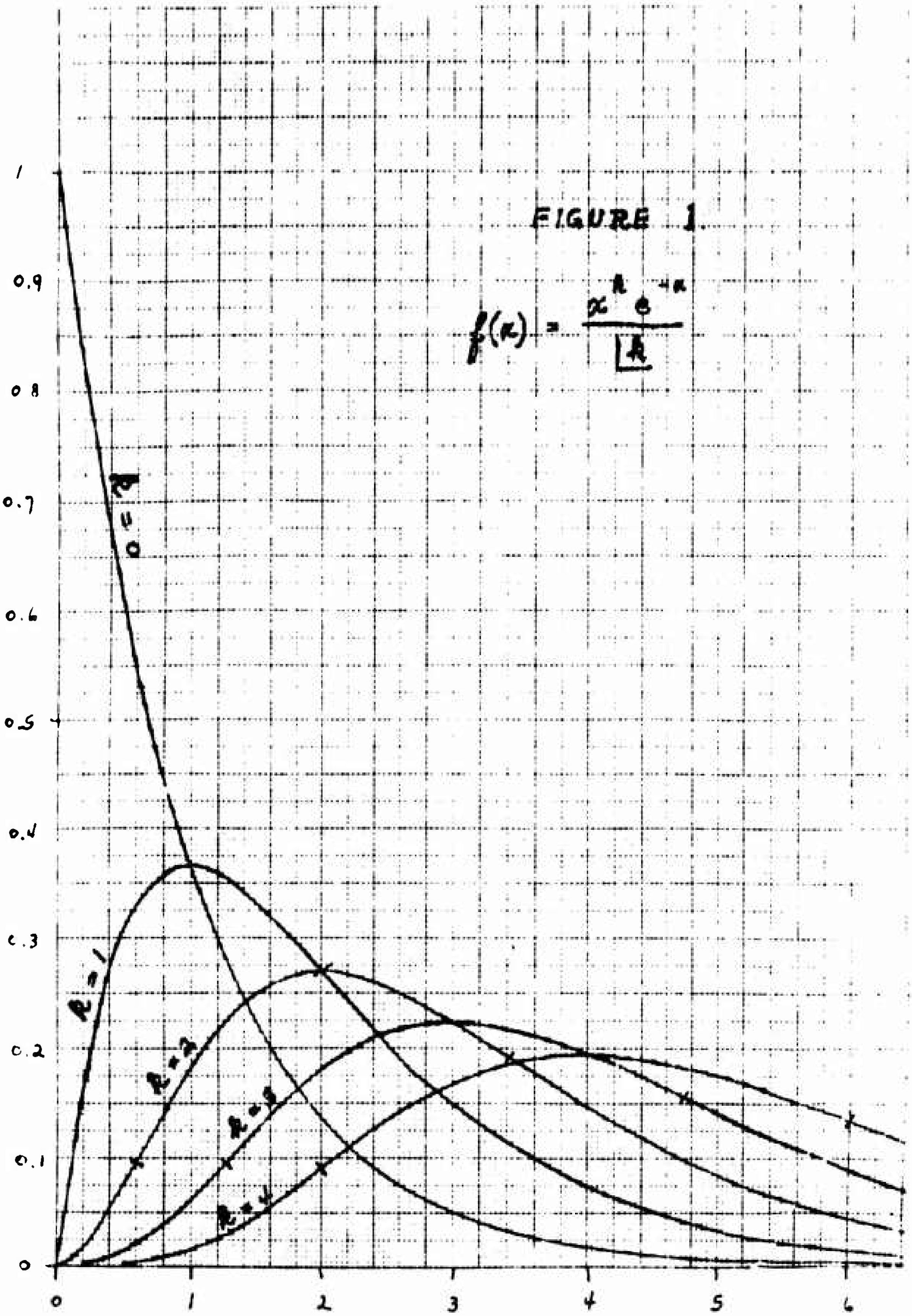


FIGURE 1
$f(x) = \frac{x^k e^{-x}}{\lfloor k}$
k = 0
k = 1
k = 2
k = 3
k = 4
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0
1
2
3
4
5
6

$$2)\quad f'(x) = \frac{1}{\underline{|k}}\left\{k x^{k-1} e^{-x} - e^{-x} x^{k}\right\} = \frac{x^{k-1} e^{-x}}{\underline{|k}}(k - x) \tag{6}$$

Thus a maximum occurs when $x = k$.

$$3)\qquad f''(x) = \frac{x^{k-2} e^{-x}}{\underline{|k}} = \left\{k(k-1) - 2kx + x^{2}\right\} \tag{7}$$

A point of inflection is found whenever

$$x^{2} - 2kx + k(k - 1) = 0,$$

i.e., when $x = k \pm \sqrt{k}$. For some programming purposes, when $k = 1$, the origin may serve as the missing point of inflection. The slope there is unity.

4) Every curve crosses every other curve exactly once, and in consecutive order.

5) Two consecutive curves intersect at the maximum point of the second, since the only non-trivial solution of

$$\frac{x^{k} e^{-x}}{\underline{|k}} = \frac{x^{k+1} e^{-x}}{\underline{|k+1}}$$

occurs when $x = k + 1$.

4. <u>TRANSFORMING THE PROBABILITY FUNCTION.</u> If the case $k = 0$ is treated separately, the transformation $w = x/k$ suggests itself. Letting $x = kw$, $dx = kdw$ and it is seen that

$$\int_{kw=0}^{\infty} \frac{(kw)^{k} e^{-kw}}{\underline{|k}}\, kdw = 1,$$

since merely employing the transformation will not affect the value of this definite integral. But the probability function <u>in w</u> is

$$g(w) = \frac{ke^{-kw} (kw)^{k}}{\underline{|k}} \tag{8}$$

Basically, this transformation rescales the abscissae by $1/k$, and hence the densities (ordinates) by k, thereby preserving area. Several graphs of this function are shown in Figure 2. Notice that every curve has its maximum point at $w = 1$. Also, $\bar{w} \to 1$. Points of inflection occur at $1 \pm 1/\sqrt{k}$.

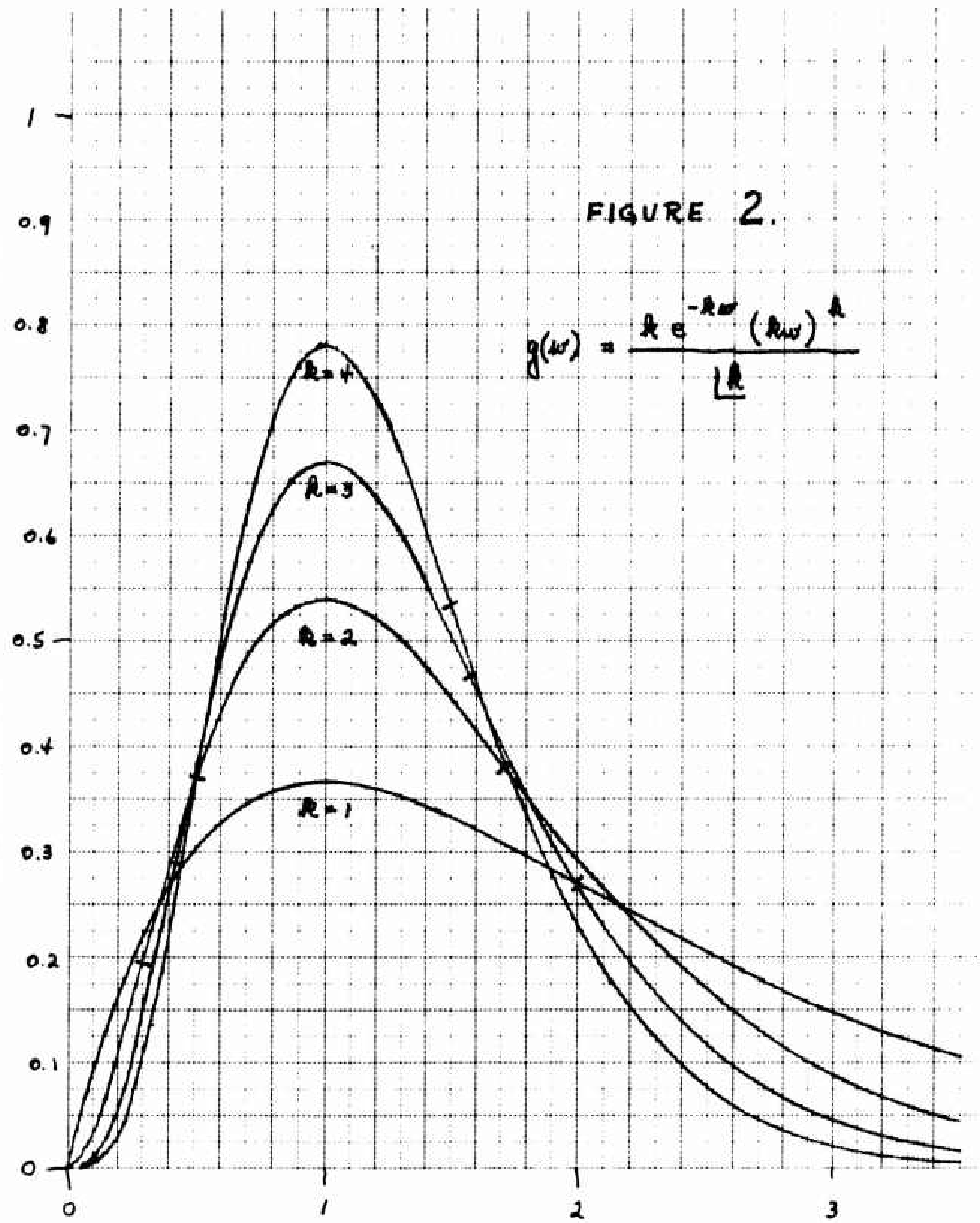

$g(w) = \dfrac{k\,e^{-kw}\,(kw)^{k}}{\lfloor k}$
k=4
k=3
k=2
k=1
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
1
2
3

FIGURE 2.

Although the transformation is useful for studying this family of functions, it matters very little whether levels of confidence are computed from

$$\int_{x=0}^{z} f(x)\, dx \quad \text{or} \quad \int_{w=0}^{z/k} g(w)\, dw.$$

In this paper, the form in $f(x)$ will be used.

5. INTEGRATION BY PARTS. When a function is defined by (or can be described as) a definite integral, very frequently it will be found that repeated integration by parts will produce an expansion suitable for computing. In fact, as in the instance at hand, it may be possible to expand in either ascending or descending factorials (or powers, as the case may be), thereby producing two different expansions, both of which are valid. Usually, one will appear in the familiar form of a power series which converges more rapidly for smaller values of the argument. The other will be the associated asymptotic expansion. If the parameter which appears in the factorial part of the probability function can be restricted to integer values only, the asymptotic expansion becomes finite in length and is an exact expression.

The sought probability integral can be stated

$$P(z) = \int_{x=0}^{z} f(x)\, dx = \int_{x=0}^{z} \frac{x^k e^{-x}}{\underline{|k}}\, dx \tag{9}$$

and gives the probability that x does not exceed the (perhaps arbitrary) value z.

Can the indefinite integral $\int \frac{x^k e^{-x}}{\underline{|k}}\, dx$ be evaluated by parts, k being a fixed, positive integer?

Let $u = e^{-x}$ and $dv = \frac{x^k}{\underline{|k}}\, dx$.

Then $du = -e^{-x}dx$ and $v = \frac{x^{k+1}}{\underline{|k+1}}$.

$$\text{Thus} \int \frac{x^k e^{-x}}{\underline{|k}}\, dx = \frac{e^{-x} x^{k+1}}{\underline{|k+1}} + \int \frac{e^{-x} x^{k+1}}{\underline{|k+1}}\, dx$$

It is apparent at once that the second integral is like the first, save k has been augmented by unity. It is clear that the process can be reapplied endlessly, yielding

$$\int \frac{x^k e^{-x}}{\underline{|k}}\, dx = \sum_{i=1}^{\infty} \frac{e^{-x} x^{k+i}}{\underline{|k+i}}, \quad (i = 1, 2, 3, \ldots).$$

Passing to the lower limit of the definite integral ($x = 0$), the sum vanishes, since x factors every term. (It may be more correct to say that the sum reduces to the constant of integration.) Thus

$$P(z) = \int_{x=0}^{z} \frac{x^k e^{-x}}{\underline{|k}}\, dx = \sum_{i=1}^{\infty} \frac{e^{-z} z^{k+i}}{\underline{|k+i}} \tag{10}$$

The term-to-term recurrence ratio is $z/(k + i)$. Since z is constant while $(k + i)$ increases without bound, the series will (eventually) converge for all positive values of z.

Now let $u = \frac{x^k}{\underline{|k}}$ and $dv = e^{-x}dx$. Then

$$du = \frac{kx^{k-1}}{\underline{|k}}\, dx = \frac{x^{k-1}}{\underline{|k-1}}\, dx \text{ and } v = -e^{-x}, \text{ whence}$$

$$\int \frac{x^k\ e^{-x}}{\underline{|k}}\, dx = -e^{-x} \frac{x^k}{\underline{|k}} + \int \frac{x^{k-1}\ e^{-x}}{\underline{|k-1}}\, dx.$$

Noting that $-e^{-x}$ will factor every term, we can write the result in the form

$$\int \frac{x^k\ e^{-x}}{\underline{|k}}\, dx = -e^{-x} \left\{ \frac{x^k}{\underline{|k}} + \frac{x^{k-1}}{\underline{|k-1}} + \ldots + \frac{x^2}{2} + x + 1 \right\}$$

At the lower limit ($x = 0$), the right-hand member becomes

$$\lim_{x \to 0} -e^{-x} \left\{ \left(\frac{x^0}{\underline{|0}} \right) \right\} = -1 \left(\frac{1}{1} \right) = -1.$$

The definite integral thus is given by

$$P(z) = \int_{x=0}^{z} \frac{x^k\ e^{-x}}{\underline{|k}}\, dx = 1 - e^{-z} \left\{ \frac{z^k}{\underline{|k}} + \frac{z^{k-1}}{\underline{|k-1}} + \ldots z + 1 \right\} \tag{11}$$

Are the two solutions equivalent? Is it true that

$$\sum_{i=1}^{\infty} \frac{e^{-z}\ z^{k+i}}{\underline{|k+i}} = 1 - e^{-z} \left\{ 1 + z + \frac{z^2}{2} + \ldots + \frac{z^{k-1}}{\underline{|k-1}} + \frac{z^k}{\underline{|k}} \right\} \quad ?$$

Multiplying by e^z and transposing, it is seen that

$$\left\{1 + z + \ldots + \frac{z^k}{\underline{|k}}\right\} + \sum_{i=1}^{\infty} \frac{z^{k+i}}{\underline{|k + i}} = e^z$$

is the well-known Maclaurin series for e^z. Therefore the two solutions are indeed equivalent.

It is a fact that if the upper limit of integration be taken at the maximum ($w = 1$; i.e., $z = k$), the level of confidence will always be less than 1/2 and hence of little statistical interest. (See Table 1.) However, the argument $z = k$ has an important use of a different sort. It enables us to select a series for computing whose terms are known to decrease monotonically. This results in worthwhile economy for larger values of k. There are two cases to consider.

First: Let $0 < z \leq k$. The series

$$P(z) = \sum_{i=1}^{\infty} \frac{e^{-z} z^{k+i}}{\underline{|k + i}} \tag{10}$$

is chosen for use. Obviously, the term-to-term recurrence ratio is given by $z/(k + i)$. Under the stated conditions, this is always less than unity.

Second: Let $z > k$. The formula

$$P(z) = 1 - e^{-z} \left\{\frac{z^k}{\underline{|k}} + \frac{z^{k-1}}{\underline{|k - 1}} + \ldots + \frac{z^2}{2} + z + 1\right\} \tag{11}$$

is used. The recurrence ratio is

$$\frac{k + 1 - i}{z}, \quad (i = 1, 2, 3, \ldots, k)$$

which again is less than unity. For large values of k, the interior series can be summed as though it were an infinite series, thus achieving a laudable saving in the number of terms required.

<u>6. COMPUTING A LEVEL OF CONFIDENCE ($z > k$).</u> The value of z may be derived from any source, or it may be arbitrarily specified. The proper formula, as we have seen, is

$$P(z) = 1 - e^{-z} \left\{\frac{z^k}{\underline{|k}} + \frac{z^{k-1}}{\underline{|k - 1}} + \ldots + z + 1\right\} \tag{11}$$

TABLE 1

CONFIDENCE LEVEL AT MAXIMUM ORDINATE

k	$\int_0^k f(x)\,dx$
0	0.000000
1	0.264241
2	0.323324
3	0.352768
4	0.371163
5	0.384039
6	0.393697
7	0.401286
8	0.407453
9	0.412592
10	0.416960
12	0.424035
15	0.431910
20	0.440907
30	0.451648
50	0.462483
100	0.473438
200	0.481206
400	0.486706
1000	0.491591

When k is small ($k < 12$, say), the resulting finite expression submits easily to direct computation. But when k is very large, two difficulties arise.

First: The number of terms becomes excessive. If the series is summed as though it were an infinite series--i.e., the relative size of each new term is observed--the process can be truncated when additional terms no longer affect the result in the computer.

Second: Large factorials will overflow the computer. To circumvent this, the first term of the series is computed by logarithms. Stirling's formula ($k > 11$) is given by

$$\ln_e \underline{|k} = 0.91893\ 85332 + \left(k + \frac{1}{2}\right) \ln_e k$$

$$- k + \frac{1}{12k} \left\{1 - \frac{1}{30k^2} \left(1 - \frac{2}{7k^2}\right)\right\} . \qquad (12)$$

The first term is (disregarding sign) $\frac{e^{-z} z^k}{\underline{|k}}$; hence, its logarithm will be $k \ln_e z - z - \ln_e \underline{|k}$, which should not cause overflow within the range of useful numbers.

7. <u>COMPUTING z WHEN A LEVEL OF CONFIDENCE IS SPECIFIED.</u> ($L = P(z) > 0.5$) No new formula is available for the inverse. Instead, successive approximations z_0, z_1, z_2, . . . are computed until a steady state is reached. Newton's method serves very well. See the discussion in [4] pp. 279-280.

For any z_i, compute $P(z_i)$, $f(z_i)$ and $f'(z_i)$. The required incremental area is of course $P(z) - P(z_i)$. We approximate this area with a trapezoid of width Δz whose ordinates are $f(z_i)$ and $f(z_i) + \Delta z f'(z_i)$. We have seen earlier that the first term of the wanted series for $P(z_i)$ is

$$\frac{e^{-z_i} z_i^k}{\underline{|k}}$$

$$\text{Also } f(z_i) = \frac{e^{-z_i} z_i^k}{\underline{|k}};$$

$$\text{and } f'(z_i) = \left\{\frac{k}{z_i} - 1\right\} f(z_i).$$

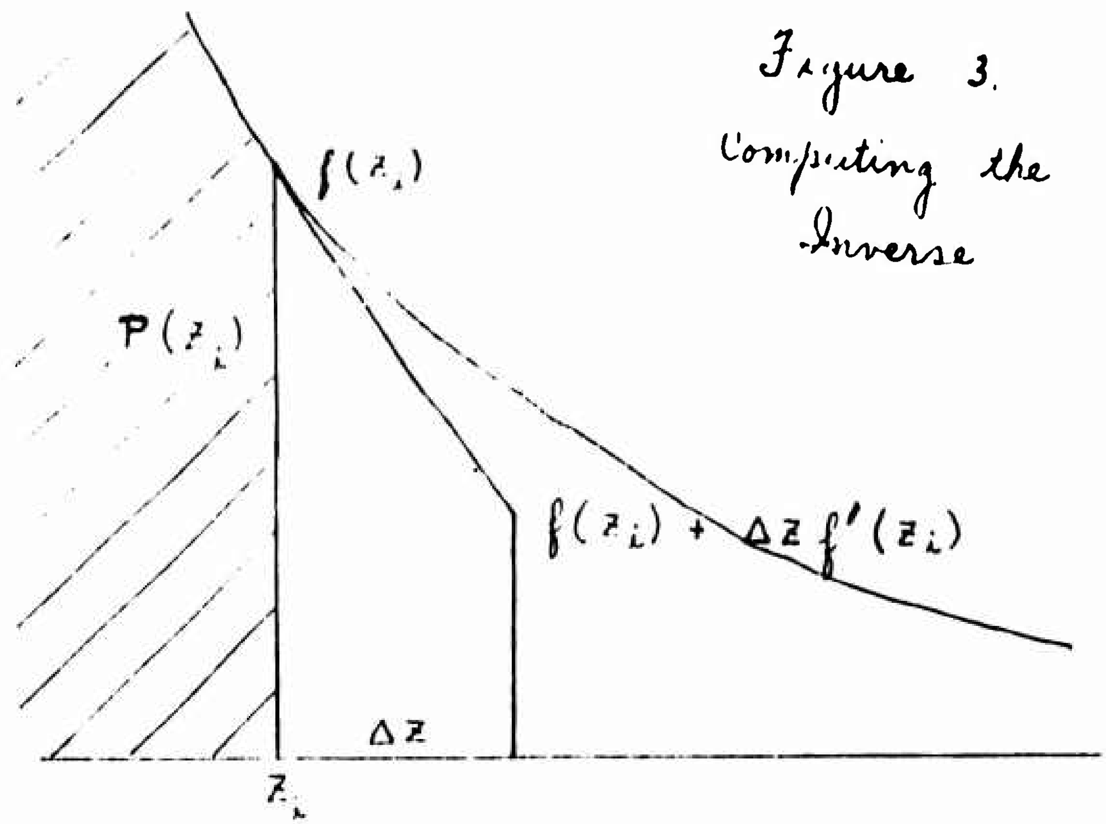


Figure 3.
Computing the Inverse

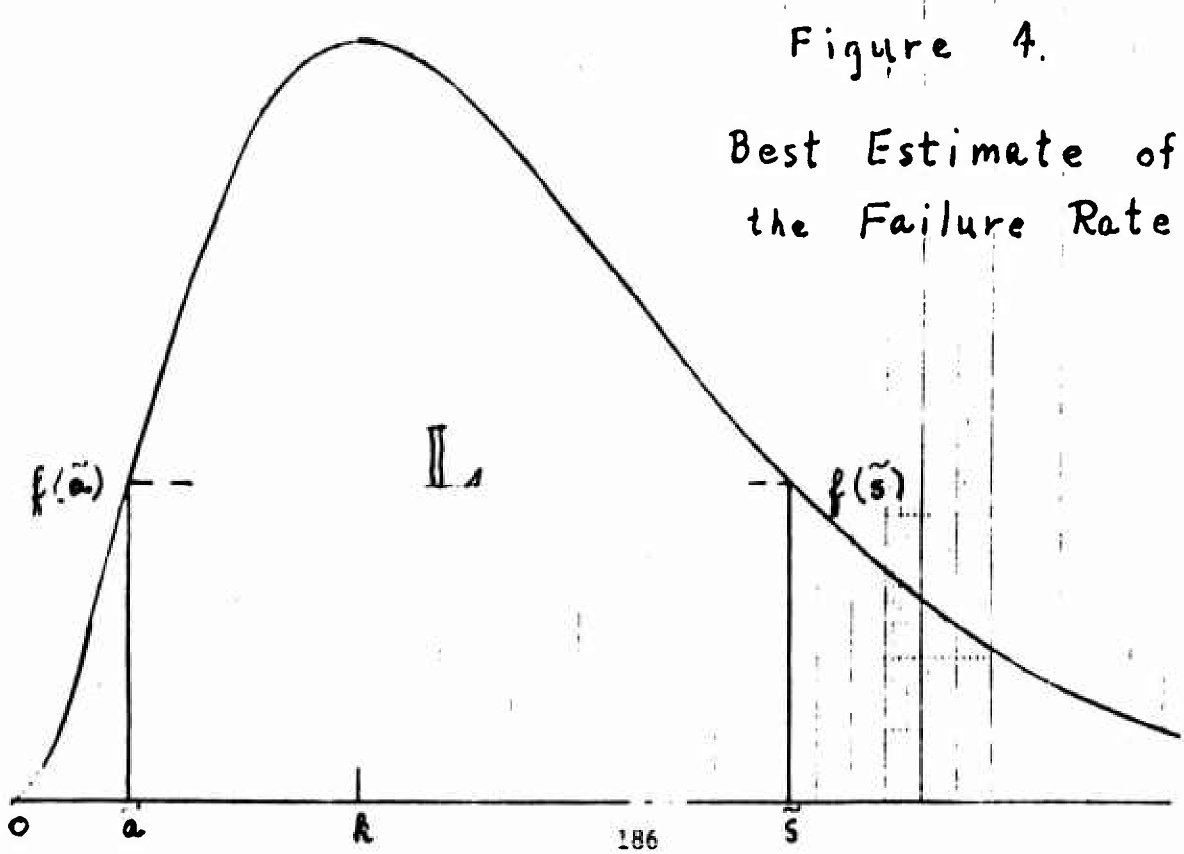


Figure 4.
Best Estimate of the Failure Rate

The approximating trapezoid is given by

$$P(z) - P(z_i) \approx \Delta z \left\{f(z_i) + \frac{\Delta z}{2} f'(z_i)\right\}$$

which can be solved for Δz.

$$\Delta z \approx \frac{-f(z_i) \pm \sqrt{[f(z_i)]^2 + 2f'(z_i)\,[P(z) - P(z_i)]}}{f'(z_i)}$$

Since ultimately $\Delta z \to 0$, it is apparent that the positive square root yields the true solution. Noting that

$$\frac{f(z_i)}{f'(z_i)} = \frac{z_i}{k - z_i}$$

the formula can be simplified to

$$\Delta z \approx \frac{z_i}{z_i - k} - \sqrt{\left(\frac{z_i}{z_i - k}\right)^2 + \frac{2[P(z) - P(z_i)]}{f'(z_i)}} \tag{13}$$

The process is stable when started from the right-hand point of inflection; i.e.,

$$z_0 = k + \sqrt{k} \tag{14}$$

8. THE BEST ESTIMATE OF THE FAILURE RATE.* For a specified level of confidence L, the general solution of the probability integral is

$$L = \int_a^s f(x)\,dx.$$

There are, of course, an unlimited number of solution pairs (a, s) which satisfy this equation. Up to this point, we have concerned ourselves with the case a = o. This form properly is used to test for compliance with an imposed standard.

Sometimes, however, that standard is absent, unrealistic, or even erroneous. But it is still required to make a meaningful statement about the failure rate. In this situation, the *Best Estimate* is recommended. Essentially, that solution pair (a, s) is chosen which minimizes the difference | s - a |.

---

*See [4] pp. 267-270.

Values of a and s thus determined are designated by a tilde ($\tilde{a}$, $\tilde{s}$).

Some properties of the Best Estimate of the Failure Rate are:

a. $\tilde{s} - \tilde{a}$ is minimum, by definition.

b. The limits of integration lie on opposite sides of the maximum; i.e., $\tilde{a} < k < \tilde{s}$.

c. The ordinates at $\tilde{a}$ and $\tilde{s}$ are equal; i.e, $f(\tilde{a}) = f(\tilde{s})$.

d. The solution is unique.

There are several steps in the solution.

Step One. For any s_i, compute $f(s_i)$, $f'(s_i)$, $P(s_i)$.

(To begin, set $s_0 = k + \sqrt{k}$.)

Step Two. For <u>each</u> s_i, solve for the value $a < k$ such that $f(a) = f(s_i)$.

For any a_j, compute $f(a_j)$ and $f'(a_j)$.

Then

$$\Delta a = \frac{f(s_i) - f(a_j)}{f'(a_j)}. \tag{15}$$

The process is repeated until f(a) and hence a is found to the desired accuracy. This value of f(a) is then associated with $f(s_i)$ by appending the subscript i. (The subscript j is dropped, being no longer necessary.) For every <u>new</u> value of s_i, the a-process is begun afresh by setting $a_j = k - \sqrt{k}$.

Step Three. The value $f(a_i) = f(s_i)$ having been found, compute $P(a_i)$. (The values for a_i and $f'(a_i)$ will already have been computed.) The desired incremental area is $L - P(s_i) + P(a_i)$

Step Four. The incremental area always will appear in two separate parts. The ratio of these areas can be estimated quite closely by the slopes. Thus

$$\left(L - P(b_i) + P(a_i)\right)\left(\frac{f'(a_i)}{f'(a_i) - f'(s_i)}\right)$$

will appear on the right. It is convenient to express the ratio in terms of the ordinates.

$$\frac{f'(a_i)}{f'(a_i) - f'(s_i)} = \frac{\left(\frac{k - a_i}{a_i}\right) f(a_i)}{\left(\frac{k - a_i}{a_i}\right) f(a_i) - \left(\frac{k - s_i}{s_i}\right) f(s_i)}$$

But since $f(a_i) = f(s_i)$, this value can be cancelled from numerator and denominator, leaving

$$\frac{f'(a_i)}{f'(a_i) - f'(s_i)} = \frac{s_i(k - a_i)}{k(s_i - a_i)}. \tag{16}$$

Thus a suitable approximating trapezoid is given by

$$\left(L - P(s_i) + P(a_i)\right) \left(\frac{s_i(k - a_i)}{k(s_i - a_i)}\right) = \Delta s \ \left\{f(s_i) + \frac{\Delta s}{2} f'(s_i)\right\} \tag{17}$$

which can be solved for Δs by the method of Section 7, above.

9. <u>EXPRESSING RESULTS IN TERMS OF MEAN LIFE.</u> It should be noted that the methods developed in this paper are virtually independent of the definition of *Event*. (*Event* often will be synonymous with *Duration of Test*.) Suitable values of $\tilde{a}$ and $\tilde{s}$ (or z, as the case may be) having been found, it is apparent that they should be expressed in the units *failures per event*. If at this point the definition of *event* is imposed, the results can be expressed in *failures per mile* or *failures per hour* or whatever.

Now the simple reciprocal converts to *mean life*. It should be remembered that taking the reciprocal reverses the sense of inequality signs.

APPENDIX A

CHI-SQUARE AND OTHER POISSON-RELATED FUNCTIONS

Let us define the following special functions:

Incomplete exponential function:

$$e_n(x) = 1 + x + \frac{x^2}{\underline{|2}} + \ldots + \frac{x^n}{\underline{|n}}.$$

The series consists of n + 1 terms.

Gamma function:

$$\Gamma(x) = \int_0^\infty e^{-t}\, t^{x-1}\, dt \qquad (x > o)$$

$$\text{Thus } \Gamma(x+1) = \int_0^\infty e^{-t}\, t^{x}\, dt.$$

Incomplete gamma function:

$$\gamma(x,z) = \int_0^z e^{-t}\, t^{x-1}\, dt \qquad (x > o)$$

and, of course, $o < z < \infty$.

Prym's function:

$$\Gamma(x,z) = \int_z^\infty e^{-t}\, t^{x-1}\, dt \qquad (z > o)$$

Immediately it is seen that

$$\gamma(x,z) + \Gamma(x,z) = \Gamma(x)$$

and that dividing both sides of this equation by $\Gamma(x)$ will produce a probability relationship.

Thus we can state

$$P(x,z) = \frac{\gamma(x,z)}{\Gamma(x)} = 1 - \frac{\Gamma(x,z)}{\Gamma(x)}.$$

Now for any particular problem, x and hence $\Gamma(x)$ will remain fixed. In terms of Prym's function we can write

$$P(x,z) = 1 - \frac{1}{\Gamma(x)} \int_z^\infty e^{-t} t^{x-1} dt.$$

It is easy to develop $\Gamma(x,z)$, using repeated integration by parts.* It is found that

$$\Gamma(x,z) \sim e^{-z} z^{x-1} \sum_{s=0}^{\infty} \frac{(x-1)(x-2)\ldots(x-s)}{z^s}$$

is a valid asymptotic expansion for fixed x and large z.

When x is an integer, the series terminates.

When x is not an integer, the terms of the series alternate in sign after $s > x$. The series diverges after $s > x + z$.

Let us replace x with k in the formulae in order that x can be employed as a variable of integration. Thus the formulae restated appear as follows:

$$P(k,z) = 1 - \frac{\Gamma(k,z)}{\Gamma(k)} = 1 - \frac{1}{e^z}\left[e_{k-1}(z)\right]$$

$$P(k,z) = 1 - \frac{e^{-z} z^{k-1}}{\Gamma(k)} \sum_{s=0}^{k-1} \frac{(k-1)(k-2)\ldots(k-s)}{z^s}$$

When k is a positive integer.

This case of k being a positive integer was studied at length in its application to sampling distributions by Helmert (1876) and K. Pearson (1900). Thus arose the statistics of the χ^2 distribution. The exponent 2 in χ^2 has little significance beyond ensuring that the parameter is non-negative.

*See [3] p. 66.

The χ^2 probability function* is defined by:

$$P(\chi^2|\nu) = \frac{1}{2^{\nu/2}\,\Gamma(\nu/2)} \int_0^{\chi^2} (t)^{(\nu/2)-1}\, e^{-t/2}\, dt$$

$$Q(\chi^2|\nu) = \frac{1}{2^{\nu/2}\,\Gamma(\nu/2)} \int_{\chi^2}^{\infty} (t)^{(\nu/2)-1}\, e^{-t/2}\, dt$$

$$P(\chi^2|\nu) + Q(\chi^2|\nu) = 1$$

Comparing this to the earlier-derived

$$P(k,z) = \frac{\gamma(k,z)}{\Gamma(k)} = \frac{1}{\Gamma(k)} \int_{x=0}^{z} e^{-x}\, x^{k-1}\, dx,$$

it is seen that the only differences are in the scaling of the parameters. For let $\nu = 2k$. Then

$$P(\chi^2|2k) = \frac{1}{2^k\Gamma(k)} \int_0^{\chi^2} t^{k-1}\, e^{-t/2}\, dt$$

$$= \frac{1}{\Gamma(k)} \int_0^{\chi^2} \left(\frac{t}{2}\right)^{k-1} \left(\frac{1}{2}\, e^{-t/2}\right) dt$$

Now let $t = 2x$, from which $dt = 2dx$. Replacing the variable of integration,

$$P(\chi^2|2k) = \frac{1}{\Gamma(x)} \int_{2x=0}^{2x=\chi^2} x^{k-1}\, e^{-x}\, dx$$

and it is seen that $\chi^2 = 2z$ properly scales the limit of integration.

When $\nu = 2k$ is an ODD integer, two things happen. $\Gamma(k)$ contains the factor $\sqrt{\pi}$ and $\sum\limits_{s=0}^{\infty} \dfrac{(k-1)\,(k-2)\,.\,.\,.\,(k-s)}{z^s}$ does not terminate. The behavior (accuracy) of the asymptotic expansion near $z \approx k - s$ must be investigated.

---

*See [1] 26.4 page 940.

APPENDIX B

PROGRAM PLANNING - POISSON

1. <u>INTRODUCTION.</u> As a general rule, the only variable of observation will be k, the number of failures. The variable of integration will be x, with z one of its extreme values (limits of integration).

It is necessary to define *event* in some suitable unit (time, distance, mass, volume, etc.); e.g., *event* = 4240 hours. *Event* often is synonymous with *Duration of Test*.

Many formulae of interest are greatly simplified if expressed as functions of f(x) or of f(z). Thus

$$f(x) = \frac{e^{-x} x^{k}}{|\underline{k}}$$

$$f'(x) = \left(\frac{k}{x} - 1\right) f(x)$$

$$f''(x) = \left(\frac{k^2 - k}{x^2} - \frac{2k}{x} + 1\right) f(x)$$

$$P(z) = \sum_{i=1}^{\infty} \frac{e^{-z} z^{k+i}}{|\underline{k+i}} \qquad (i = 1, 2, 3, \ldots)$$

$$= T_1 + T_2 + T_3 + \ldots + T_i + \ldots$$

$$T_1 = \frac{z}{k+1} f(z) \text{ and}$$

$$T_j = \frac{z}{k+j} T_{j-1}$$

Also, $1 - P(z) = T_0 + T_1 + T_2 + \ldots + T_i + \ldots$

$$T_0 = f(z)$$

$$T_{j+1} = \frac{k - j}{z} T_j$$

This latter series terminates when $k = j$.

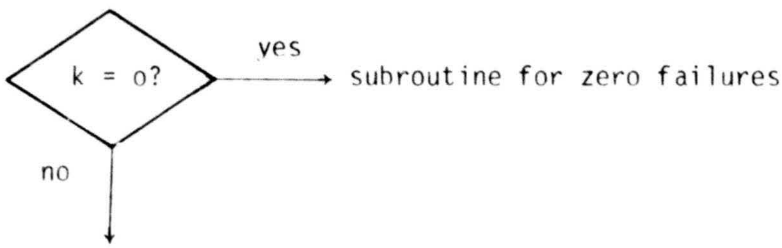
For large values of z, compute f(z) by logarithms.

$$\ln_e f(z) = k \ln_e z - z - \ln_e \underline{|k}$$

Stirling's formula for $\ln_e \underline{|k}$ is useful here. If k does not change, it need be computed but once.

<u>2. COMPUTING L (z specified)</u>. Equations (4), (5), (9), (10) and (11).

Enter data



Compute $\ln_e \underline{|k}$. (If $k > 15$, use Stirling's formula.)

Compute $k \ln_e z - z$.

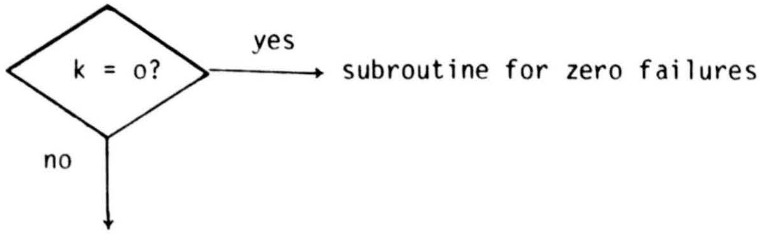
Compute $f(z) = \dfrac{z^k e^{-z}}{\underline{|k}} = T_0$

Compute $L = \int_0^z f(x)\, dx$ from one of the methods in the previous paragraph.

$\left(\text{If } k > 15 \text{ } \underline{\text{and}} \text{ } z < k, \text{ use } L = \sum_{i=1}^{\infty} \dfrac{e^{-z} z^{k+i}}{\underline{|k + i}} \right)$

3. COMPUTING z (L specified). To the above, add equations (6), (7) and (13).

Enter data



Subsequent portion of method assumes $L > \frac{1}{2}$.

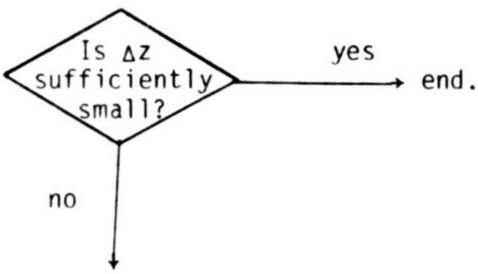
Assign $z_0 = k + \sqrt{k}$

Label $\boxed{1}$

Apply method of paragraph 2 above to compute $L_0 = \int_0^{z_0} f(x)\,dx$

Compute $f'(z_0) = \left(\frac{k}{z_0} - 1\right) f(z_0)$ and

$$\Delta z \approx \frac{z_i}{z_{i-k}} - \sqrt{\left(\frac{z_i}{z_i - k}\right)^2 + \frac{2(L - L_i)}{f'(z_i)}}$$



Assign $z_{i+1} = z_i + \Delta z$

Return to Label $\boxed{1}$.

4. COMPUTING *BEST ESTIMATE OF THE FAILURE RATE* (L specified.)

Enter data

$k = o?$ — yes → Employ paragraph 3, above.

no ↓

Assign $s_o = k + \sqrt{k}$

Label $\boxed{2}$

Compute $\int_o^{s_i} f(x)\,dx$ by method of paragraph 2 above.

Compute $f'(s_i) = \left(\frac{k}{s_i} - 1\right) f(s_i)$

$k = 1?$ — yes → assign $a_o = f(s_i)$

no → assign $a_o = k - \sqrt{k}$

Label $\boxed{3}$

Compute $f(a_j)$ and $f'(a_j)$

Compute $\Delta a = \dfrac{f(s_i) - f(a_j)}{f'(a_j)}$

Compute and store $a_{j+1} = a_j + \Delta a$

Is Δa sufficiently small? — no → Return to Label $\boxed{3}$

yes ↓

Compute $\int_0^{a_i} f(x)\,dx$ by methods of paragraph 2 above.

$$\left(\text{If } k > 15,\ \int_0^{a} f(x)\,dx = \sum_{i=1}^{\infty} \frac{e^{-a}\,a^{k+i}}{|\underline{k + i}}\right)$$

The needed increment of area is

$$A = L - \int_0^{s_i} f(x)\,dx + \int_0^{a_i} f(x)\,dx$$

The approximating trapezoid yields (momentarily dropping subscripts for convenience)

$$\Delta s = \frac{s}{k - s}\left\{1 - \sqrt{1 + \frac{2(k - a)\,(k - s)}{k(s - a)} \cdot \frac{A}{f(s)}}\right\}$$

Is Δs sufficiently small? — yes → end.

no ↓

$s_{i+1} = s_i + \Delta s$

Return to Label $\boxed{2}$.

BIBLIOGRAPHY

[1] Abramowitz, M. and Stegun, I. A., Eds., HANDBOOK OF MATHEMATICAL FUNCTIONS; Dover Publications, Inc., New York, 1972.

[2] Boole, G., THE CALCULUS OF FINITE DIFFERENCES; 5th Ed., Chelsea Publishing Co., New York, 1970.

[3] Olver, F. W. J., ASYMPTOTICS AND SPECIAL FUNCTIONS; Academic Press, New York, 1974.

[4] Rankin, D. W., ESTIMATING RELIABILITY FROM SMALL SAMPLES; in Proc. 22nd Conf. on D.O.E. in Army R.D.&T., Department of Defense, 1977.

[5] Whittaker, E. T. and Watson G. N., A COURSE OF MODERN ANALYSIS; 4th Ed., Cambridge University Press, 1963.